

Fluctuation relations and coarse-graining

Saar Rahav¹, Christopher Jarzynski^{1,2}

¹ Department of Chemistry and Biochemistry, University of Maryland, College Park 20742, USA.

² Institute for Physical Science and Technology, University of Maryland, College Park 20742, USA.

Abstract. We consider the application of fluctuation relations to the dynamics of coarse-grained systems, as might arise in a hypothetical experiment in which a system is monitored with a low-resolution measuring apparatus. We analyze a stochastic, Markovian jump process with a specific structure that lends itself naturally to coarse-graining. A perturbative analysis yields a reduced stochastic jump process that approximates the coarse-grained dynamics of the original system. This leads to a non-trivial fluctuation relation that is approximately satisfied by the coarse-grained dynamics. We illustrate our results by computing the large deviations of a particular stochastic jump process. Our results highlight the possibility that observed deviations from fluctuation relations might be due to the presence of unobserved degrees of freedom.

1. Introduction

Coarse-graining is at the heart of equilibrium statistical mechanics. It is used to describe systems with a macroscopic number of degrees of freedom with the help of a much smaller number of variables. A familiar example is the description of a gas in terms of its slowly varying local density rather than the position of all its molecules [1]. Coarse-graining is particularly important in the study of phase transitions, and more generally, interacting systems. It is a crucial building block in renormalization group theory [2, 3]. Even far from a phase transition the combination of coarse-graining with the identification of relevant and irrelevant variables can sometimes be used to replace seemingly complicated interactions by (approximate) simpler ones.

The theory of systems out of equilibrium is less developed and, in particular, the possible uses of coarse-graining in such a theory are not as clear. One approach, recently reviewed by Derrida [4], uses large deviation theory to address the fluctuations of a local density or some other coarse-grained field. This leads to a functional of the local density that characterizes the steady state and fluctuations around it.

In recent years, a different approach has been employed to study systems far from equilibrium. In this approach one investigates symmetries, related to time-reversal, that are associated with the probabilities of observing sequences of events. Such investigations have led to the discovery of a number of interesting results, collectively known as *fluctuation relations*, including some that provide information related to the likelihood to observe so-called “second law violations” [5, 6, 7]. Loosely speaking, fluctuation relations have the form

$$\ln \left[\frac{P(\Delta S = qT)}{P(\Delta S = -qT)} \right] = qT + o(T), \quad (1)$$

where $P(\Delta S)$ is the probability to observe a change ΔS in entropy, T is the observation time, and q is the observed entropy creation rate. [The notation $o(T)$ indicates a correction that grows more slowly than T , that is $\lim_{T \rightarrow \infty} o(T)/T = 0$.] These results are quite general, they remain valid far from thermal equilibrium, and they are part of an active field of research focused on the application of the laws of thermodynamics to microscopic systems [8].

Fluctuation relations are derived by considering statistical distributions of microscopic trajectories of the system of interest. For any such trajectory, we assign a value to a quantity ΔS , which we interpret physically as a measure of entropy production. By comparing the probability of observing a given trajectory (γ) with that of its time-reversed counterpart ($\bar{\gamma}$), the symmetry represented by equation (1) emerges (see section 3). This approach, however, runs into difficulties when coarse-graining is involved. Consider an experiment in which some degrees of freedom cannot be observed, perhaps because they evolve on time scales much faster than the response time of the detector. In this case our observation of the system gives us a *coarse-grained* – i.e. a “smeared”, or locally averaged – trajectory, rather than a full microscopic record of the evolving state of the system. It is not at all evident that such coarse-grained trajectories satisfy the same symmetry relations as their fully microscopic counterparts. Moreover, important information might be lost due to coarse graining; specifically, the expression for entropy production, ΔS , might contain contributions from those degrees of freedom that are not resolved by the detector. Under these circumstances, can we define some sort of “coarse-grained entropy production”, ΔS^{CG} , which satisfies a fluctuation relation?

Our goal in this paper is to give partial answers to such questions. We will examine a simple system for which an assumed separation of time scales leads to a simple coarse-graining transformation. The system we study is a stochastic jump process whose probability distribution satisfies a Master equation. Fluctuation relations for such jump processes were first derived by Lebowitz and Spohn [9], following Kurchan's analysis of Langevin processes [10]. These two papers laid the groundwork for much subsequent work on stochastic fluctuation relations, which has recently been reviewed by Harris and Schütz [11]. In the present paper, we impose a separation of time scales and argue that, when the slow component of the dynamics is accurately measured, a non-trivial fluctuation relation, which differs from the one valid for the microscopic system, is (approximately) satisfied. This relation involves a coarse-grained entropy production, which depends only on the slow component of the evolution.

We note that in the context of fluctuation theorems, an explicit separation of time scales has appeared as well in the work of Zamponi *et.al.* [12]. These authors have studied a Langevin particle interacting with a nonequilibrium thermal environment consisting of slow and fast components at different temperatures (see also [13]). For such a system a fluctuation-dissipation relation between spontaneous and externally induced fluctuations can be used to define an *effective temperature*. The system was shown to satisfy a generalized fluctuation theorem formulated in terms of this effective temperature.

In section 2 we will present the model used and apply a coarse-graining transformation, replacing it by a simpler system. The (approximate) coarse-grained system is also Markovian and therefore satisfies a fluctuation relation of its own. In section 3 we use the existence of a coarse-grained counterpart to define coarse-grained trajectories, and a coarse-grained entropy production ΔS^{CG} . These trajectories can be heuristically viewed as measurements obtained with a limited-resolution detector. We show that the coarse-grained entropy production approximately satisfies a fluctuation relation (for the original system). A simple example is numerically studied in section 4, in order to illuminate the results obtained in previous sections. The results are summarized in section 5.

2. Stochastic systems with simple coarse-graining

While there are many physically relevant and interesting examples of systems with natural coarse-grained counterparts, we will focus on stochastic jump processes governed by master equations. After briefly reviewing general properties of such processes, we present models with special structure, which makes them amenable to coarse-graining. We then use simple perturbation theory to motivate a coarse-graining transformation that replaces the system of interest by a smaller, reduced system.

Consider a system composed of a finite number of distinct states. The system can make transitions between these states, and these transitions are assumed to be described by a Markov process. Specifically, given two states σ and σ' , the rate $R(\sigma|\sigma')$ is the probability per unit time to make a transition to σ , from a current state σ' . While not all transitions are allowed (some of the rates vanish), we do assume that if the transition from σ' to σ is allowed [$R(\sigma|\sigma') \neq 0$] then also the reverse transition is allowed [$R(\sigma'|\sigma) \neq 0$]. This assumption is analogous, but not equivalent, to microscopic reversibility for deterministic systems.

It is useful to depict such a system with the help of a graph, in which each node represents a state and each link represents a pair of non-vanishing transitions. For

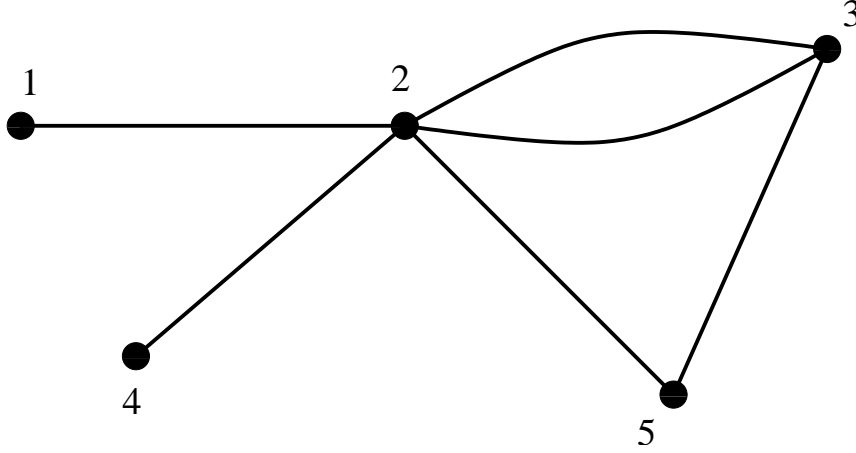


Figure 1. A stochastic system described with the help of a graph. The nodes correspond to the states of the system. Each link represents a non-vanishing transition between two states, and its reversed counterpart. The system is fully specified when the transition rates are known.

instance, the graph shown in figure 1 illustrates a five-state system with six non-vanishing links. Note that we allow multiple links between a given pair of states, as between states 2 and 3. In other words, there may exist information that allows one to distinguish between different ways to make a transition. This implies that four rates – namely, $R_{\alpha=1,2}(3|2)$ and $R_{\alpha=1,2}(2|3)$ – are needed to describe all transitions between states 2 and 3.

Such a process satisfies a master equation

$$\frac{dP(\sigma, t)}{dt} = \sum_{\sigma'} R(\sigma|\sigma') P(\sigma', t), \quad (2)$$

where $P(\sigma, t)$ denotes the normalized probability to find the system in the state σ at time t . It is useful to think of $P(t)$ as a vector whose components are labeled by the index σ , and \mathbf{R} as a matrix. A non-diagonal element $R(\sigma|\sigma')$ is given by the sum of the rates from σ' to σ , e.g. $R(\sigma|\sigma') = \sum_{\alpha} R_{\alpha}(\sigma|\sigma')$, whereas a diagonal element $R(\sigma|\sigma)$ specifies the net probability rate of transitions out of the state σ . To ensure conservation of probability, the rates satisfy

$$\sum_{\sigma} R(\sigma|\sigma') = 0. \quad (3)$$

We assume that the graph is *connected*, that is, it is possible to go from any state to any other state using the links. In that case, the solution of equation (2) decays with time to a unique stationary solution P^s satisfying [14]

$$\sum_{\sigma'} R(\sigma|\sigma') P^s(\sigma') = 0. \quad (4)$$

The stationary probability current $J_{\alpha}^s(\sigma', \sigma) = R_{\alpha}(\sigma|\sigma') P^s(\sigma') - R_{\alpha}(\sigma'|\sigma) P^s(\sigma)$ measures the net flow from σ' to σ , via the transition α , in the stationary state. If $J_{\alpha}^s = 0$ for every transition, then *detailed balance* is satisfied, and it is natural to think of the system as being in a state of thermal equilibrium. For a generic transition

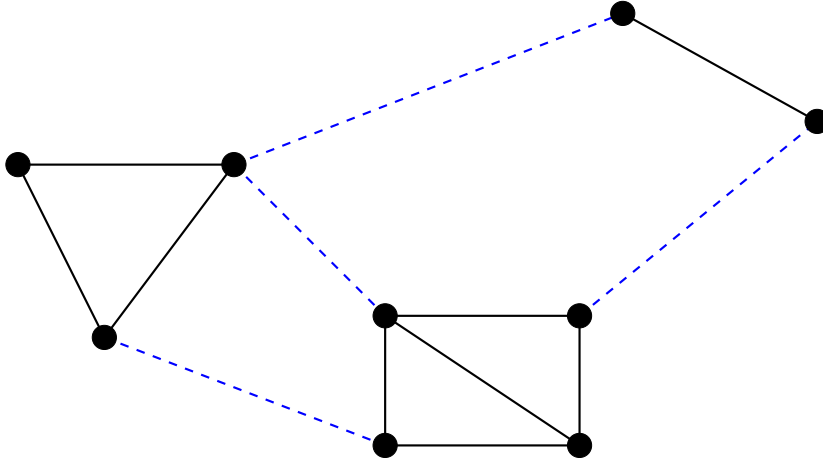


Figure 2. Graph of a stochastic system with a meaningful coarse-graining transformation. The solid lines denote links with large transition rates (of order unity). The dotted lines denote weakly connected states (with rates of order ε).

matrix \mathbf{R} , however, the currents J_α^s are non-vanishing, and then we consider the system to be in an out-of-equilibrium steady state [15].

Systems described by master equations are, generally, relatively easy to handle mathematically. However, our interest is in systems for which it is useful to coarse-grain the dynamics. This leads us to examine systems with special structure. Specifically, we will assume that the N states of our system are assorted into L clusters, with “strong” links within a given cluster, $R(\sigma|\sigma') \sim 1$, and “weak” links between states belonging to different clusters, $R(\sigma|\sigma') \sim \varepsilon \ll 1$, as illustrated in figure 2 for $N = 9$, $L = 3$. We further assume that each separate cluster is (internally) connected, and also that the entire system is connected. Finally, we assume that there is at most one microscopic transition connecting any pair of states in the system[‡]. (It will become clear that coarse-graining does not preserve this property.) A system satisfying these conditions is characterized by two widely separated time scales, corresponding to the fast dynamics of relaxation within each cluster ($\tau_f \sim 1$), and the slow redistribution of probability among the clusters ($\tau_s \sim \varepsilon^{-1}$).

The microscopic evolution of our system is fully described by a sequence of transitions from state to state, $\sigma \rightarrow \sigma' \rightarrow \sigma'' \dots$, along with the times at which these transitions occur. We introduce coarse graining by imagining that we monitor this evolution using a low-resolution apparatus, capable of distinguishing different clusters, but not different states within a cluster. The dynamics we observe then consists of transitions from cluster to cluster, $c \rightarrow c' \rightarrow c'' \dots$, which provides only partial information regarding the underlying microscopic trajectory. In general this low-resolution dynamics is not a Markov process, and cannot be described by a master equation of the form given by equation (2). We aim to investigate whether this coarse-grained sequence satisfies a meaningful fluctuation theorem.

Let us imagine, for a moment, that such a coarse-graining transformation maps the original system onto a smaller system, whose states correspond to the clusters of

[‡] The discussion can be generalized to several possible transitions easily. We use this assumption since we find the idea of a microscopic level, where the dynamics is simpler, appealing.



Figure 3. The structure of possible simple systems, obtained by coarse-graining the system depicted in figure 2.

the original system. For instance, the system depicted in figure 2 is mapped onto a 3-state system. Note that there may be several different transitions between a pair of clusters. The ability of the low-resolution apparatus to distinguish between these transitions determines the structure of the simpler system. For instance, the system depicted in figure 2 will be mapped to the one depicted in figure 3a when the apparatus can distinguish between the transitions connecting the same clusters. Alternatively, when the apparatus can only distinguish the clusters, but can not be used to identify the transitions, the system will be mapped to the one depicted in figure 3b.

In order to give a more quantitative meaning to the notion of coarse-graining, we now develop a perturbation theory that maps our original N -state stochastic jump process onto a *reduced* L -state stochastic jump process; each state of the reduced system represents one of the clusters in the original system. Such a mapping is useful only if it preserves relevant dynamical properties. In our case this means the reduced dynamics ought to provide a good approximation of the low-resolution dynamics of the original system. In what follows we show that the structure that we have assumed above leads to a separation between fast and slow decay modes of equation (2). The latter are then used to set the transition rates of the reduced, L -state process, in such a way that this process does, indeed, faithfully reproduce the low-resolution dynamics of the original system, to leading order in ε .

For systems with the cluster structure depicted in figure 2, we decompose the transition matrix, \mathbf{R} :

$$R(\sigma|\sigma') = R^{(0)}(\sigma|\sigma') + \varepsilon R^{(1)}(\sigma|\sigma'). \quad (5)$$

Here $\mathbf{R}^{(0)}$ is block-diagonal, with each block corresponding to one cluster [see equation (31) for an example]. The non-diagonal elements of $\varepsilon\mathbf{R}^{(1)}$ are the weak links connecting states belonging to different clusters, while the diagonal elements of $\varepsilon\mathbf{R}^{(1)}$ are adjusted so that equation (3) holds separately for $\mathbf{R}^{(0)}$ and $\mathbf{R}^{(1)}$. We assume that the dependence of \mathbf{R} on ε resides entirely in the factor that multiplies $\mathbf{R}^{(1)}$ in the above decomposition. Thus all non-vanishing elements of $\mathbf{R}^{(0)}$ and $\mathbf{R}^{(1)}$ are assumed to be order unity. We then treat $\varepsilon\mathbf{R}^{(1)}$ as a small perturbation to $\mathbf{R}^{(0)}$.

To keep considerations simple, let us assume that both \mathbf{R} and $\mathbf{R}^{(0)}$ have full sets of left and right eigenvectors, satisfying bi-orthogonality. Thus for \mathbf{R} we have

$$\begin{aligned} \mathbf{R}\mathbf{v}_n &= \Lambda_n \mathbf{v}_n \\ \mathbf{w}_n^T \mathbf{R} &= \Lambda_n \mathbf{w}_n^T \\ \mathbf{w}_m^T \cdot \mathbf{v}_n &= \delta_{mn}, \end{aligned} \quad (6)$$

with $m, n = 1, 2, \dots, N$. The same relations hold among the eigenvalues and the left and right eigenvectors of $\mathbf{R}^{(0)}$, denoted by $\Lambda_n^{(0)}$, $\mathbf{w}_n^{(0)}$ and $\mathbf{v}_n^{(0)}$.

The transition matrix \mathbf{R} has one vanishing eigenvalue, $\Lambda_1 = 0$. All other eigenvalues are negative. The corresponding right eigenvector \mathbf{v}_1 is proportional to the steady state distribution, and the left eigenvector \mathbf{w}_1 expresses conservation of probability. Since bi-orthogonality does not fix the normalization of these eigenvectors, we choose, for convenience, $\mathbf{v}_1 = \mathbf{P}^s$ and $\mathbf{w}_1^T = (1, 1, \dots, 1)$.

To carry out our perturbation analysis, we expand eigenvectors and eigenvalues in powers of ε ,

$$\mathbf{v}_n = \sum_{l=0}^{\infty} \varepsilon^l \mathbf{v}_n^{(l)} \quad , \quad \mathbf{w}_n = \sum_{l=0}^{\infty} \varepsilon^l \mathbf{w}_n^{(l)} \quad , \quad \Lambda_n = \sum_{l=0}^{\infty} \varepsilon^l \Lambda_n^{(l)} \quad , \quad (7)$$

and we match powers of ε in equation (6). At leading order we get

$$\mathbf{R}^{(0)} \mathbf{v}_n^{(0)} = \Lambda_n^{(0)} \mathbf{v}_n^{(0)}, \quad (8)$$

and similarly for $\mathbf{w}_n^{(0)}$. Making use of the block-diagonal structure of $\mathbf{R}^{(0)}$, and diagonalizing each block separately, we obtain eigenvectors $\mathbf{v}_n^{(0)}$ and $\mathbf{w}_n^{(0)}$ that inherit this structure: each eigenvector has non-vanishing elements only in one block, corresponding to one cluster.

The block-diagonal structure of $\mathbf{R}^{(0)}$ describes L mutually isolated clusters, each of which supports its own steady state and corresponding vanishing eigenvalue. Thus $\mathbf{R}^{(0)}$ has L vanishing eigenvalues, $\Lambda_n^{(0)} = 0$ for $n = 1, \dots, L$. The corresponding right eigenvectors, $\mathbf{v}_n^{(0)}$, are composed of blocks of vanishing elements with the steady state of the n 'th cluster, $\mathfrak{P}^s(n; \sigma)$ inserted at the appropriate place. (Here σ runs over states belonging to the n 'th cluster.) The corresponding left eigenvectors, $\mathbf{w}_n^{(0)}$, have a similar structure, with the elements $\mathfrak{P}^s(n; \sigma)$ replaced by 1's. The remaining $N - L$ eigenvectors of $\mathbf{R}^{(0)}$ have negative eigenvalues of order unity.

In first-order perturbation theory, the term $\varepsilon \mathbf{R}^{(1)}$ in equation (5) shifts each eigenvalue by an amount of order ε . While this represents a relatively small change in the non-vanishing eigenvalues $\Lambda_{L+1}^{(0)}, \dots, \Lambda_N^{(0)}$, it crucially lifts the degeneracy among the first L (vanishing) eigenvalues. These are replaced by a single vanishing eigenvalue (corresponding to the steady state of the fully connected system) and $L - 1$ negative eigenvalues of order ε , which describe the long time decay of the system. Assuming for simplicity that the degeneracy is fully lifted at first order, higher order corrections will lead only to small further changes in the eigenvalues and eigenvectors. Therefore, in the following we will concentrate on the application of first-order perturbation theory to the degenerate sector, $n \leq L$.

Degenerate perturbation theory begins with the $L \times L$ matrix

$$\Pi(n|m) \equiv \mathbf{w}_n^{(0)T} \mathbf{R}^{(1)} \mathbf{v}_m^{(0)} \quad , \quad m, n = 1, \dots, L, \quad (9)$$

whose eigenvalues represent the first-order corrections, $\Lambda_n^{(1)}$, to the vanishing eigenvalues of $\mathbf{R}^{(0)}$. Because of the correspondence between the clusters of the system and the left/right eigenvectors in the above equation, the quantity $\Pi(n|m)$ represents the transition rate between clusters m and n . In the following we emphasize this identification by using indices $c, c' = 1, \dots, L$ to specify the clusters.

The slow decay modes of the system are determined by the matrix Π whose dimensionality is the number of clusters, L . This matrix satisfies

$$\sum_{c=1}^L \Pi(c|c') = \sum_{c=1}^L \sum_{\sigma, \sigma'=1}^N w_c^{(0)}(\sigma) R^{(1)}(\sigma|\sigma') v_{c'}^{(0)}(\sigma') = \sum_{\sigma, \sigma'=1}^N R^{(1)}(\sigma|\sigma') v_{c'}^{(0)}(\sigma') = 0, \quad (10)$$

where we have used $\sum_{c=1}^L w_c^{(0)}(\sigma) = 1$. In other words Π satisfies the conservation-of-probability condition expressed (for the full system) by equation (3). It is thus meaningful to interpret the equation,

$$\frac{d\tilde{P}(c, t)}{dt} = \sum_{c'} \varepsilon \Pi(c|c') \tilde{P}(c', t), \quad (11)$$

as describing a Markov jump process. We now discuss the physical meaning that can be assigned to Π and $\tilde{P}(c, t)$.

The separation of time scales between the $L - 1$ slow and the $N - L$ fast decay modes, implies that after the decay of the latter the probability distribution $P(\sigma, t)$ satisfies $P(\sigma, t) \simeq \mathfrak{P}^s(c; \sigma) \tilde{P}(c, t)$. Thus the probability inside each cluster is a product of its local steady state distribution, \mathfrak{P}^s , and the total probability to find the system in the cluster, $\tilde{P}(c, t) = \sum_{\sigma \in c} P(\sigma, t)$. Equation (11) approximately describes the time evolution of the latter. More precisely, equation (11) governs a *reduced* (L -state) Markov jump process, and *this reduced process provides an approximate description of the transitions between clusters within the original system*. Note that the individual elements of Π differ from those of $\mathbf{R}^{(1)}$: for a transition out of a cluster c , with the rate ε_+ appearing in $\varepsilon \mathbf{R}^{(1)}$, the probability current is $\varepsilon_+ P(\sigma, t) \simeq \varepsilon_+ \mathfrak{P}^s(c; \sigma) \tilde{P}(c, t)$. This contributes a rate $\varepsilon_+ \mathfrak{P}^s(c; \sigma)$ in the appropriate place in $\varepsilon \Pi$.

The procedure outlined above is not the only way to define a reduced system. An alternative approach would involve measuring the transition rates between clusters in the steady state of the full system, and then constructing a reduced, L -state Markov jump process governed by precisely these transition rates. We expect that this “empirical construction” – based on the observed cluster-to-cluster transition rates – and the “perturbative construction” leading to equation (11), will give reduced systems that become equivalent in the limit $\varepsilon \rightarrow 0$.

To summarize, we have used perturbation theory to give quantitative meaning to a coarse-graining transformation, in which a Markov jump process among N states, arranged into L clusters, is replaced by a reduced Markov jump process among L states meant to represent those clusters. In the following section, where we consider fluctuations, it is important to bear in mind that the reduced dynamics only approximately reproduce the statistics of transitions between clusters of the original system. Of course, one might improve the approximation by extending the perturbation expansion to higher orders of ε ; see, e.g. the Appendix of reference [16]. However, higher-order corrections to the eigenvectors of a particular degenerate sector typically involve *all* unperturbed eigenvectors, not only those within that sector. Thus it is not generally possible to reduce the dimensionality of the system when considering higher orders in ε .

3. Fluctuation relations for systems that can be coarse-grained

In this section we ask whether the structure described in section 2 leads to a non-trivial fluctuation relation. We start by briefly reviewing the fluctuation relation for general

stochastic jump processes [9, 17, 18, 11]. We will then turn to systems amenable to coarse-graining.

Fluctuation relations are obtained by comparing probability densities of microscopic paths and their time reversed counterparts. A path, or a history, $\gamma(T)$, will denote a stochastic trajectory, represented by an ordered sequence of transitions at given times (between say 0 and T),

$$\gamma \equiv \sigma_0 \xrightarrow{t_1, \alpha_1} \sigma_1 \xrightarrow{t_2, \alpha_2} \sigma_2 \longrightarrow \dots \xrightarrow{t_n, \alpha_n} \sigma_n. \quad (12)$$

The α 's in equation (12) denote the transitions between states. They are redundant for the original system, but would be needed after coarse-graining, where several transitions between clusters are possible. We have chosen to explicitly denote the transitions in order to use similar notation for the original and coarse-grained system. The conditional probability density (or weight) corresponding to a path is

$$w(\gamma) = \prod_{i=0}^n \exp[R(\sigma_i|\sigma_i)(t_{i+1} - t_i)] \prod_{j=1}^n R_{\alpha_j}(\sigma_j|\sigma_{j-1}). \quad (13)$$

In equation (13) $-R(\sigma_i|\sigma_i)$ denotes the overall rate of probability flow out of the state σ_i , and we set $t_0 = 0$ and $t_{n+1} = T$. The probability density of a path is obtained by multiplying $w(\gamma)$ by the probability to sample its initial conditions, $P(\sigma_0, t = 0)$.

To each path we can define a time reversed counterpart $\bar{\gamma}(t) \equiv \gamma(T - t)$, starting at σ_n at $t = 0$, and ending at σ_0 at $t = T$. Since the path γ and its counterpart $\bar{\gamma}$ reside in the same states for the same lengths of time their exponential survival probability factors in equation (13) are identical. As a result

$$\frac{\mathcal{P}(\gamma)}{\mathcal{P}(\bar{\gamma})} = \frac{P(\sigma_0, t = 0)w(\gamma)}{P(\sigma_n, t = 0)w(\bar{\gamma})} = \frac{P(\sigma_0, t = 0)}{P(\sigma_n, t = 0)} \prod_{i=1}^n \frac{R_{\alpha_i}(\sigma_i|\sigma_{i-1})}{R_{\alpha_i}(\sigma_{i-1}|\sigma_i)}. \quad (14)$$

Taking the logarithm of both sides yields

$$\ln \frac{\mathcal{P}(\gamma)}{\mathcal{P}(\bar{\gamma})} = \ln \frac{P(\sigma_0, t = 0)}{P(\sigma_n, t = 0)} + \sum_{i=1}^n \ln \frac{R_{\alpha_i}(\sigma_i|\sigma_{i-1})}{R_{\alpha_i}(\sigma_{i-1}|\sigma_i)} \equiv B + \sum_{i=1}^n \delta S_i. \quad (15)$$

The right side includes a boundary term, B , arising from the initial and final states, and a sum of contributions from the n transitions. If we interpret

$$\delta S = \ln \frac{R_{\alpha}(\sigma|\sigma')}{R_{\alpha}(\sigma'|\sigma)} \quad (16)$$

as the amount of entropy that is produced when the system makes a transition α from σ' to σ [9, 17, 11], then

$$\Delta S = \sum_{i=1}^n \delta S_i \quad (17)$$

is the net entropy production associated with the trajectory γ , and we have

$$\ln \frac{\mathcal{P}(\gamma)}{\mathcal{P}(\bar{\gamma})} = B + \Delta S. \quad (18)$$

Identities of the form of equation (18) are at the heart of both *steady-state* and *transient* fluctuation relations [11]. The distinction between the two cases, roughly, amounts to whether one chooses to consider the limit $T \rightarrow \infty$ (in which case the relative contribution of B can be neglected, at least for finite systems [9]), or rather to argue that the boundary term B represents the net change in the entropy of the

system (in which case the sum on the right represents the entropy produced in the system and its surroundings [17]). In either situation the desired fluctuation relation follows in a more or less straightforward manner.

For our purposes the distinction between the transient and steady state cases is not of central importance. Rather, we aim to investigate whether a relationship similar to equation (18) can be obtained for systems of the type studied in section 2. To clarify this issue, let us return to the situation introduced briefly in section 2, in which we monitor the evolution of our system using a low-resolution apparatus that distinguishes between clusters, but not between states within a cluster. Under these circumstances we do not observe the full trajectory, γ , but rather a coarse-grained version Γ (defined below) that describes only transitions between clusters. Can we expect this trajectory to satisfy a fluctuation relation? In what follows we use the perturbation analysis of section 2 to define a coarse-grained entropy production, ΔS^{CG} , that differs from ΔS above, and is expressed in terms of the *observed* transition rates. We argue that this quantity approximately satisfies a fluctuation relation, and that the approximation improves as the separation of time scales becomes more pronounced (i.e. as $\varepsilon \rightarrow 0$). The deviations from this (approximate) relation are more pronounced for extreme values of ΔS^{CG} , which select rare trajectories that, in some sense, do not respect the separation of time scales. These deviations reflect the non-Markovian nature of the observed, coarse-grained dynamics, and can be viewed as evidence of the existence of unobserved degrees of freedom.

The reduced, L -state system defined in section 2 is a stochastic, Markov jump process that satisfies an exact fluctuation relation. The entropy produced during a transition α from c to c' is given in this case by

$$\delta S^{\text{red}} = \ln \frac{\Pi_\alpha(c|c')}{\Pi_\alpha(c'|c)}, \quad (19)$$

and the net entropy production associated with a trajectory is just a sum of such contributions. To simplify notations, we assume here, and in what follows, the scenario where different transitions between the same clusters are distinguishable. The following considerations can be easily modified to apply to coarse-grained dynamics with indistinguishable transitions.

Let us now return to the full, N -state system, but assume the conditions of a low-resolution apparatus. Suppose the evolution of the system from $t = 0$ to $t = T$ is described by a trajectory γ . To this trajectory let us assign a *coarse-grained* entropy production $\Delta S^{\text{CG}}[\gamma]$, which is a sum of contributions from jumps between different clusters. Specifically, when the system makes a transition α from cluster c' to cluster c , ΔS^{CG} is updated by an amount

$$\delta S^{\text{CG}} = \ln \frac{\Pi_\alpha(c|c')}{\Pi_\alpha(c'|c)}. \quad (20)$$

There are no changes to ΔS^{CG} arising from (unobserved) jumps within clusters. Thus equation (20) defines a coarse-grained entropy production for the N -state system, under low resolution, by borrowing the definition that arises in the reduced, L -state Markov system, equation (19). We now argue that

$$\Delta S^{\text{CG}} = \sum_k' \delta S_k^{\text{CG}} \quad (21)$$

(the sum includes only cluster-to-cluster transitions) satisfies an approximate fluctuation relation.

Since only the transitions between different clusters are observed under low-resolution conditions, let us define a coarse-grained path

$$\Gamma \equiv \left\{ \gamma | c_0 \xrightarrow{\tau_1, \alpha_1} c_1 \xrightarrow{\tau_2, \alpha_2} \dots c_{n'} \right\}. \quad (22)$$

We interpret Γ as the collection of all paths γ , of duration T , that start in the cluster c_0 at time 0, make the transition α_1 to cluster c_1 at time τ_1 and so forth. Thus while the inter-cluster transitions of all trajectories in Γ are identical, the intra-cluster transitions typically differ among these trajectories. Note that we retain information on the identity (α) of a transition between two clusters, if only to identify the microscopic states involved. Moreover, since ΔS^{CG} has the same value for every $\gamma \in \Gamma$, we take this value to define the coarse-grained entropy production associated with Γ .

The probability density assigned to the coarse-grained path Γ is

$$\mathcal{P}(\Gamma) = \sum_{\gamma \in \Gamma} \mathcal{P}(\gamma). \quad (23)$$

This is computed by summing over all possible combinations of transitions inside the cluster and integrating over their corresponding transition times, keeping the inter-cluster transitions fixed. Each coarse-grained trajectory has a time reversed counterpart $\bar{\Gamma}$, which passes through the clusters in the reversed order, at the appropriate times. Let us now compare the statistical weights associated with such a *conjugate pair* of coarse-grained trajectories, Γ and $\bar{\Gamma}$.

The weight of a coarse-grained trajectory is a product of factors (partial weights) resulting from the different segments of the microscopic trajectories. A *segment* here refers to the interval of time – bracketed by a pair of transition times, say τ_k and τ_{k+1} – during which the system remains in exactly one cluster. [E.g. equation (22) specifies a path with $n' + 1$ segments.]

To compute the partial weight of a given segment, consider the system shown in figure 4, with five states and two outgoing links. The latter represent the decay of the probability to remain within this system: when a trajectory makes a transition along such a link, it departs the system and does not return. Let $\mu(\sigma, t | \sigma', t_0)$ denote the conditional probability to find the system at (σ, t) , given that it was at (σ', t_0) .

The partial weights discussed above are contributions from trajectory segments that perform exactly this type of stochastic dynamics. By specifying Γ one specifies all the transitions between the clusters. Given such a transition, α_k , let us denote the microscopic states before and after the transition by σ_k^- and σ_k^+ , respectively. The contribution from all trajectory segments entering c_k at (σ_k^+, τ_k) and leaving this cluster for the first time at $(\sigma_{k+1}^-, \tau_{k+1})$, is just $\mu(\sigma_{k+1}^-, \tau_{k+1} | \sigma_k^+, \tau_k)$. §

For the systems studied here, the separation of time scales suggests that the system is likely to relax rapidly within a cluster before departing that cluster, which in turn suggests a simple approximation:

$$\mu(\sigma_{k+1}^-, \tau_{k+1} | \sigma_k^+, \tau_k) \simeq \exp[\varepsilon \Pi(c_k, c_k)(\tau_{k+1} - \tau_k)] \mathfrak{P}^s(c_k; \sigma_{k+1}^-). \quad (24)$$

Here $-\varepsilon \Pi(c_k, c_k)$ is the approximate rate of escape from c_k , assuming the system has relaxed within this cluster, while $\mathfrak{P}^s(c_k; \sigma_{k+1}^-)$ is the steady state probability to be at σ_{k+1}^- in the isolated cluster c_k . At this level of approximation, the weight μ does not

§ Note that each cluster has its own conditional probability distribution μ . We will not introduce new notation to explicitly differentiate between the different functions μ , since the identity of the cluster is already specified in the microscopic states that appear as arguments of μ .

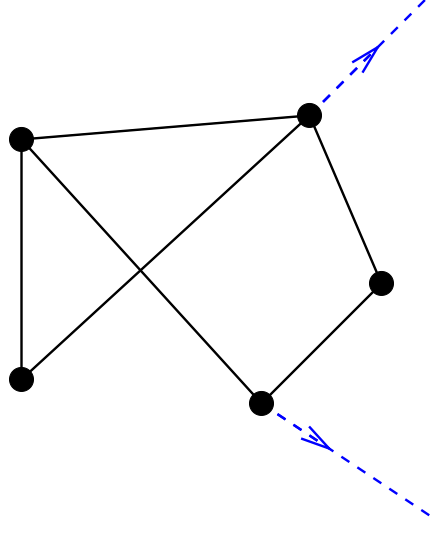


Figure 4. A cluster of states with some additional outgoing transitions. Stochastic motion in such a system is related to partial contributions to weights of coarse-grained trajectories, see text.

depend on the initial state σ_k^+ . We defer discussion of possible corrections to equation (24) to a later stage. The probability density of a coarse-grained trajectory is then

$$\mathcal{P}(\Gamma) = \left[\sum_{\sigma_0 \in c_0} P(\sigma_0, t=0) \mu(\sigma_1^-, \tau_1 | \sigma_0, 0) \right] \prod_{k=1}^{n'-1} \mu(\sigma_{k+1}^-, \tau_{k+1} | \sigma_k^+, \tau_k) \quad (25)$$

$$\times \left[\sum_{\sigma \in c_{n'}} \mu(\sigma, T | \sigma_{n'}^+, \tau_{n'}) \right] \prod_{k=1}^{n'} \varepsilon R_{\alpha_k}^{(1)}(\sigma_k^+ | \sigma_k^-).$$

The two factors in square brackets, corresponding to the initial and final clusters, account for the fact that the initial and final microstates are not fully specified by Γ . Substituting equation (24) into equation (25) leads to

$$\begin{aligned} \mathcal{P}(\Gamma) &\simeq \tilde{P}(c_0, t=0) \prod_{k=0}^{n'} \exp[\varepsilon \Pi(c_k | c_k)(\tau_{k+1} - \tau_k)] \prod_{k=1}^{n'} \left[\varepsilon R_{\alpha_k}^{(1)}(\sigma_k^+ | \sigma_k^-) \mathfrak{P}^s(c_k; \sigma_k^-) \right] \\ &= \tilde{P}(c_0, t=0) \prod_{k=0}^{n'} \exp[\varepsilon \Pi(c_k | c_k)(\tau_{k+1} - \tau_k)] \prod_{k=1}^{n'} \varepsilon \Pi_{\alpha_k}(c_k | c_{k-1}), \end{aligned} \quad (26)$$

where $\Pi_{\alpha_k}(c_k | c_{k-1})$ is the contribution of the process α_k , to the element (c_k, c_{k-1}) of the reduced transition matrix Π (see section 2). The right side of this equation is exactly the weight of a trajectory in the reduced, Markovian system whose time evolution obeys equation (11). In other words, equation (26) states that the coarse-grained trajectories Γ are described by approximately the same statistics as those of the reduced, L -state Markovian system. Evaluating the ratio of weights for a conjugate pair Γ and $\bar{\Gamma}$, and applying our definition of ΔS^{CG} [equation (21)], we get

$$\frac{\mathcal{P}(\Gamma)}{\mathcal{P}(\bar{\Gamma})} \simeq \frac{\tilde{P}(c_0, t=0)}{\tilde{P}(c_{n'}, t=0)} \cdot \exp(\Delta S^{\text{CG}}[\Gamma]). \quad (27)$$

(The factors containing the diagonal elements of Π , appearing in equation (26), cancel in this ratio.) At this point we have arrived at an analogue of equation (18), the crucial relation from which both transient and steady-state fluctuation relations are obtained. Our analysis thus reveals that the coarse-grained entropy production ΔS^{CG} – which is defined for the full, N -state system under conditions of low resolution, but which was motivated by a comparison with the reduced, L -state system of section 2 – satisfies an approximate fluctuation relation. We will rederive this result, using different considerations, in the next section.

Let us briefly take a closer look at the coarse-grained entropy, equation (20). Comparing the definitions of δS^{CG} and δS , we find

$$\delta S^{\text{CG}}(\alpha) = \delta S(\alpha) + \ln \frac{\mathfrak{P}^s(c'; \sigma^-)}{\mathfrak{P}^s(c; \sigma^+)}. \quad (28)$$

Entropy changes in stochastic jump processes are often separated into two contributions with different physical interpretations [15]. The first term on the right hand side of equation (28) expresses the change of the entropy of a thermal medium that is in contact with the system [9, 19, 18, 11, 15, 20, 21]. The second term is related to changes in the entropy of the system itself, if we interpret $-\ln P(\sigma(t), t)$ as the entropy of the system at time t along a stochastic trajectory $\sigma(t)$ [17]. Equation (28) indicates that coarse-graining does not respect the distinction between the microscopic medium and system entropies. Both contribute to changes in the coarse-grained entropy.

Note also that transitions within the clusters contribute to ΔS [equation (17)], but not to ΔS^{CG} [equation (21)]. In fact, there may be situations where the full, N -state system relaxes to a steady state characterized by non-zero stationary currents within the clusters, but transitions between clusters satisfy detailed balance. This would correspond to a state with a non-vanishing rate of entropy production ΔS , but a vanishing rate of coarse-grained entropy production, ΔS^{CG} .

The claims made in previous paragraphs rely on equation (24), which is an approximation. We should therefore consider possible corrections to equation (24) and their effects on the fluctuation relation. We will use heuristic rather than rigorous arguments to estimate these corrections.

The corrections to equation (24) depend on the time difference $\delta\tau_k \equiv \tau_{k+1} - \tau_k$. For typical time differences, comparable to the lifetime in the cluster, equation (24) is a good approximation, but we can expect it to have errors of order ε . For very short ($\delta\tau_k \sim 1$) and very long ($\delta\tau_k \sim \ln \varepsilon / \varepsilon$) time differences we may expect large errors. However, the probability of such time differences is itself of order ε . Such events are not typical.

The (coarse-grained) steady state fluctuation relation applies to the long time behavior of $(\varepsilon T)^{-1} \ln [P(\Delta S^{\text{CG}} = q\varepsilon T) / P(\Delta S^{\text{CG}} = -q\varepsilon T)]$. The only source of errors will then appear in sums over terms of the form $\ln \mu(\sigma, \tau | \sigma', \tau') / \mu(\sigma', T - \tau' | \sigma, T - \tau)$. The number of terms in the sum is proportional to the number of transitions between clusters, which in turn scales as εT . Let us consider first the corrections for a typical trajectory. The corrections are a combination of typical terms with small errors and rare terms with a large error, as discussed in the previous paragraph. The resulting errors in estimating the entropy production is small, and scales as $\varepsilon^2 T$.

The steady state fluctuation relation does not deal only with typical trajectories. In fact, one examines trajectories whose entropy production rate is (approximately) fixed. This will pick out trajectories which may be very unlikely. We *assume* that for

a finite range of entropy production rates the relevant trajectories are such that the time differences for which the approximation made in equation (24) break down are still unlikely (with probability which scales as ε). With this assumption, the coarse grained-entropy satisfies

$$\lim_{T \rightarrow \infty} \frac{1}{\varepsilon T} \ln \frac{P(\Delta S^{\text{CG}} = \varepsilon q T)}{P(\Delta S^{\text{CG}} = -\varepsilon q T)} = q + O(\varepsilon), \quad (29)$$

for some finite range of q values^{||}. Note that one can not expect this relation to hold for any coarse-grained entropy production rate q . Taking this rate to be "large enough" will eventually pick trajectories whose typical time differences are of order τ_f . These trajectories will never spend enough time in clusters to equilibrate.

4. Illustrative Example

In this section we use a simple example to illustrate fluctuation relations in a coarse-grained setting, and to gain some intuition for the admittedly abstract considerations of the previous section. We want to use the simplest system possible whose reduced counterpart can exhibit out-of-equilibrium steady states. This implies at least three clusters of states.

Consider a system with eight states, $N = 8$, organized into three clusters, $L = 3$. The transition matrix describing the full system is

$$\mathbf{R} = \mathbf{R}^{(0)} + \varepsilon \mathbf{R}^{(1)}, \quad (30)$$

with

$$\mathbf{R}^{(0)} = \begin{pmatrix} -\mathbf{3} & \mathbf{3} & \mathbf{2} & 0 & 0 & 0 & 0 & 0 \\ \mathbf{2} & -\mathbf{8} & \mathbf{3} & 0 & 0 & 0 & 0 & 0 \\ \mathbf{1} & \mathbf{5} & -\mathbf{5} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\mathbf{2} & \mathbf{3} & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{2} & -\mathbf{3} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\mathbf{4} & \mathbf{2} & \mathbf{2} \\ 0 & 0 & 0 & 0 & 0 & \mathbf{3} & -\mathbf{3} & \mathbf{6} \\ 0 & 0 & 0 & 0 & 0 & \mathbf{1} & \mathbf{1} & -\mathbf{8} \end{pmatrix}, \quad (31)$$

and

$$\mathbf{R}^{(1)} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\mathbf{4} & 0 & \mathbf{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & -\mathbf{1} & 0 & 0 & 0 & \mathbf{1} & 0 \\ 0 & \mathbf{4} & 0 & -\mathbf{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\mathbf{5} & \mathbf{3} & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{5} & -\mathbf{3} & 0 & 0 \\ 0 & 0 & \mathbf{1} & 0 & 0 & 0 & -\mathbf{1} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \quad (32)$$

(The non-vanishing rates are in bold-face.) Numbering the states of the full system from 1 to 8, the three clusters are: (1, 2, 3), (4, 5), and (6, 7, 8). We see that \mathbf{R} has the structure described in section 2: all off-diagonal elements of $\mathbf{R}^{(0)}$ are between states within a given cluster, and all off-diagonal elements of $\mathbf{R}^{(1)}$ are between states

^{||} The numerical results presented in section 4 hint that the error term in equation (29) depends on q in a non-trivial manner.

belonging to different clusters. We begin our analysis of this system by constructing the reduced, three-state Markov system.

The steady state distributions for the (unconnected) clusters appearing in $\mathbf{R}^{(0)}$ are given by the null vectors of the three matrices along the block diagonal: $\mathfrak{P}^s(1; \sigma \in 1) = \frac{1}{56}(25, 13, 18)$, $\mathfrak{P}^s(2; \sigma \in 2) = \frac{1}{5}(3, 2)$ and $\mathfrak{P}^s(3; \sigma \in 3) = \frac{1}{9}(3, 5, 1)$. From these we easily construct the eigenvectors spanning the null space of $\mathbf{R}^{(0)}$. Using equation (9) we then obtain

$$\mathbf{\Pi} = \begin{pmatrix} -\frac{5}{4} & \frac{6}{5} & \frac{5}{9} \\ \frac{13}{14} & -\frac{16}{5} & 1 \\ \frac{9}{28} & 2 & -\frac{14}{9} \end{pmatrix}. \quad (33)$$

This transition matrix defines the dynamics of the three-state reduced system, and approximately describes the cluster-to-cluster evolution of the full, eight-state system.

To evaluate the probability distribution of entropy production rates in the limit of large observation time, T , direct simulations are of limited use, as the asymptotic probability to observe a specific *average* entropy production rate decays exponentially with T . Thus obtaining good statistics becomes problematic. Following Lebowitz and Spohn [9] (see also [18]), we consider the large deviation function

$$f(q) \equiv \lim_{T \rightarrow \infty} \frac{1}{\varepsilon T} \ln P(\Delta S^{\text{CG}} = q\varepsilon T, T), \quad (34)$$

and its Legendre transform $g(\lambda)$. Here q denotes the observed entropy production rate, while $P(\Delta S^{\text{CG}})$ is the probability distribution of this entropy production. To differentiate between the large deviation function of the reduced and coarse-grained system we will denote the former (latter) by a superscript *red* (*CG*).

Lebowitz and Spohn have shown that $g(\lambda)$ is the largest eigenvalue of an operator $\mathcal{H}(\lambda)$ which governs the time evolution of averages of the type $\langle e^{-\lambda \Delta S} \rangle$. The diagonal elements of \mathcal{H} are those of \mathbf{R} , while non-diagonal elements are multiplied by a factor related to $e^{-\lambda \Delta S}$ where the entropy change is that of the transition corresponding to this non-diagonal element. We refer for references [9] and [18] for the derivation.

For the reduced system, whose transition matrix is given by equation (33), this operator is given by

$$\tilde{\mathcal{H}}(\lambda) = \varepsilon \begin{pmatrix} -\frac{5}{4} & \frac{6}{5} \left(\frac{65}{84}\right)^\lambda & \frac{5}{9} \left(\frac{81}{140}\right)^\lambda \\ \frac{13}{14} \left(\frac{84}{65}\right)^\lambda & -\frac{16}{5} & 2^\lambda \\ \frac{9}{28} \left(\frac{140}{81}\right)^\lambda & 2^{1-\lambda} & -\frac{14}{9} \end{pmatrix}. \quad (35)$$

This operator has the property $\tilde{\mathcal{H}}(\lambda) = \tilde{\mathcal{H}}^T(1 - \lambda)$, which implies the steady state (or Gallavotti-Cohen) fluctuation relation $f^{\text{red}}(q) - f^{\text{red}}(-q) = q$ [7, 9]. Note that the dependence of $\tilde{\mathcal{H}}$ on ε is trivial. To cancel this dependence we define $g^{\text{red}}(\lambda)$ to be the maximal eigenvalue of $\tilde{\mathcal{H}}$ divided by ε .

Let us turn back to the original system. We have defined a coarse-grained entropy which changes only during stochastic transitions connecting different clusters, see equation (20). One can apply the formalism developed by Lebowitz and Spohn to study the distribution function of entropy production for the coarse-grained dynamics. The derivation follows closely that of [18]. Again, the Legendre transform of the large deviation function is given by the largest eigenvalue of an operator (divided by ε).

The relevant operator is

$$\mathcal{H}(\lambda) = \begin{pmatrix} -3 & 3 & 2 & 0 & 0 & 0 & 0 & 0 \\ 2 & -8 - 4\varepsilon & 3 & 2\varepsilon \left(\frac{65}{84}\right)^\lambda & 0 & 0 & 0 & 0 \\ 1 & 5 & -5 - \varepsilon & 0 & 0 & 0 & \varepsilon \left(\frac{81}{140}\right)^\lambda & 0 \\ 0 & 4\varepsilon \left(\frac{84}{65}\right)^\lambda & 0 & -2 - 2\varepsilon & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & -3 - 5\varepsilon & 3\varepsilon 2^\lambda & 0 & 0 \\ 0 & 0 & 0 & 0 & 5\varepsilon 2^{-\lambda} & -4 - 3\varepsilon & 2 & 2 \\ 0 & 0 & \varepsilon \left(\frac{140}{81}\right)^\lambda & 0 & 0 & 3 & -3 - \varepsilon & 6 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & -8 \end{pmatrix}. \quad (36)$$

Note that only elements which connect different clusters depend on λ . This results from the fact that the coarse-grained entropy changes only during these transitions. It is important to notice that $\mathcal{H}(\lambda) \neq \mathcal{H}^T(1 - \lambda)$. This means that the coarse-grained entropy does not satisfy an *exact* fluctuation relation. It does, however satisfy an approximate one.

Consider \mathcal{H} for small values of ε . When $\varepsilon = 0$ we find that $\mathcal{H} = \mathbf{R}^{(0)}$. We have already studied the eigenvalues and eigenvectors of $\mathbf{R}^{(0)}$ in section 2, see also the discussion leading to equation (33). The largest eigenvalue is 0 with three corresponding pairs of left and right eigenvectors. All other eigenvalues are negative and of order unity. For small ε one can study the eigenvalues of $\mathcal{H}(\lambda)$ using degenerate perturbation theory. The largest eigenvalue must be part of the degenerate sector. Moreover, performing the leading order perturbation theory leads to $\tilde{\mathcal{H}}(\lambda)$ as the matrix whose eigenvalues are the leading order approximation for the eigenvalues of $\mathcal{H}(\lambda)$ that belong to the degenerate sector.

Therefore, the function $g^{\text{CG}}(\lambda)$, obtained by computing the largest eigenvalue of (36) approximately satisfies $g^{\text{CG}}(\lambda) = g^{\text{CG}}(1 - \lambda)$ (with corrections of order ε). This holds not only for the largest eigenvalue but to all the eigenvalues in the degenerate sector. This perturbation theory for the Legendre transform of the large deviation function can be considered as another justification for the existence of an (approximate) fluctuation relation for the coarse-grained entropy.

We have calculated numerically the largest eigenvalue of $\mathcal{H}(\lambda)$, as a function of λ and compared it to the largest eigenvalue obtained from $\tilde{\mathcal{H}}(\lambda)$. The results are depicted in figure 5. It is clear that the eigenvalue corresponding to the coarse-grained system is a good approximation to the one obtained for the reduced system. Moreover, the approximation improves for smaller ε . The function $g^{\text{red}}(\lambda)$ for the reduced system vanishes at $\lambda = 0$ and has the exact symmetry $g^{\text{red}}(\lambda) = g^{\text{red}}(1 - \lambda)$ as expected. The corresponding functions obtained for the coarse-grained system start to deviate from that function for larger values of $|\lambda|$. This is related to the fact that the perturbation theory cannot be uniformly valid. For instance, examining equation (36), it is clear that, for any ε , one can find λ^* so that, say, $2^{\lambda^*} \varepsilon \simeq 1$. As a result, terms which were assumed to be a small perturbation cease to be small. This can be loosely interpreted as picking contributions from highly unlikely trajectories, which may break our assumption of local relaxation.

The Legendre transform of the function $g(\lambda)$ is the large deviation function $f(q)$ which describes the asymptotic (with time) behavior of the probability distribution of the entropy production. The Legendre transform can be computed numerically. For both systems studied here the maximum of $f(q)$ was found to be at $\bar{q} \simeq 0.08$. This is the most likely value of the coarse-grained entropy production rate in a long enough

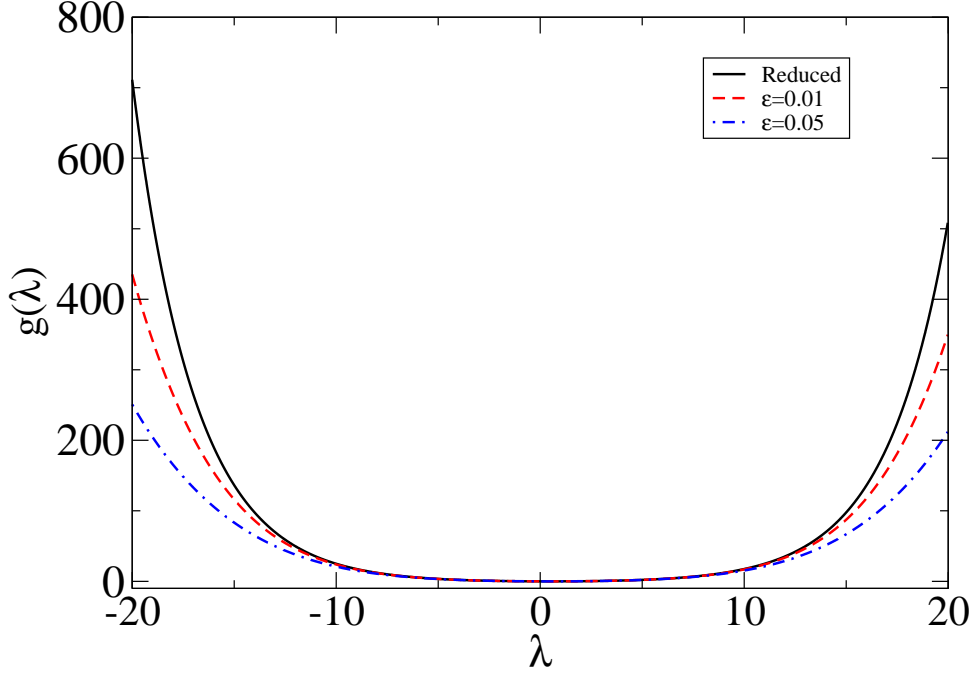


Figure 5. The largest eigenvalue of i) $\tilde{\mathcal{H}}(\lambda)$, corresponding to the reduced system (solid line), ii) $\mathcal{H}(\lambda)$, pertaining to the coarse-grained dynamics, with $\varepsilon = 0.01$ (dashed), and with $\varepsilon = 0.05$ (dashed-dotted). (All eigenvalues were divided by ε .)

experiment, in agreement with the entropy production rate calculated for the steady state of the reduced system, $\bar{q}^{\text{red}} \simeq 0.0833$. The fact that this value does not vanish indicates that the system relaxes to non-equilibrium steady state, in which detailed balance is violated.

To verify that the coarse-grained entropy (approximately) follow the Gallavotti-Cohen fluctuation relation we have calculated $f^{\text{CG}}(q) - f^{\text{CG}}(-q)$ and compared it to the expected linear behavior. The results are depicted in figure 6. It is clear that fluctuation relation is a good approximation for the coarse-grained entropy. The approximation improves for smaller values of ε , as expected. Note that the approximation deteriorates for larger values of entropy production rates. One may speculate that this is a result of the increased weight of trajectories that do not spend enough time in the clusters to relax to the local stationary state.

The numerical results, presented in this section, were for a coarse-grained entropy defined using a reduced system, which in turn was obtained from the perturbation theory of section 2. We have mentioned that there is another way to define a reduced system, with the help of the steady state of the microscopic system. The rates of this empirically-constructed reduced system differ from the ones used so far by higher orders of ε . The different rates will result in a different definition of the coarse-grained entropy. One may wonder whether this coarse-grained entropy will satisfy the fluctuation relation with much smaller deviations. We have repeated the numerical calculation performed in this section for this definition of coarse-grained entropy (with $\varepsilon = 0.05$, results not shown). The deviation from the fluctuation relation were found to be similar to those presented in figure 6.

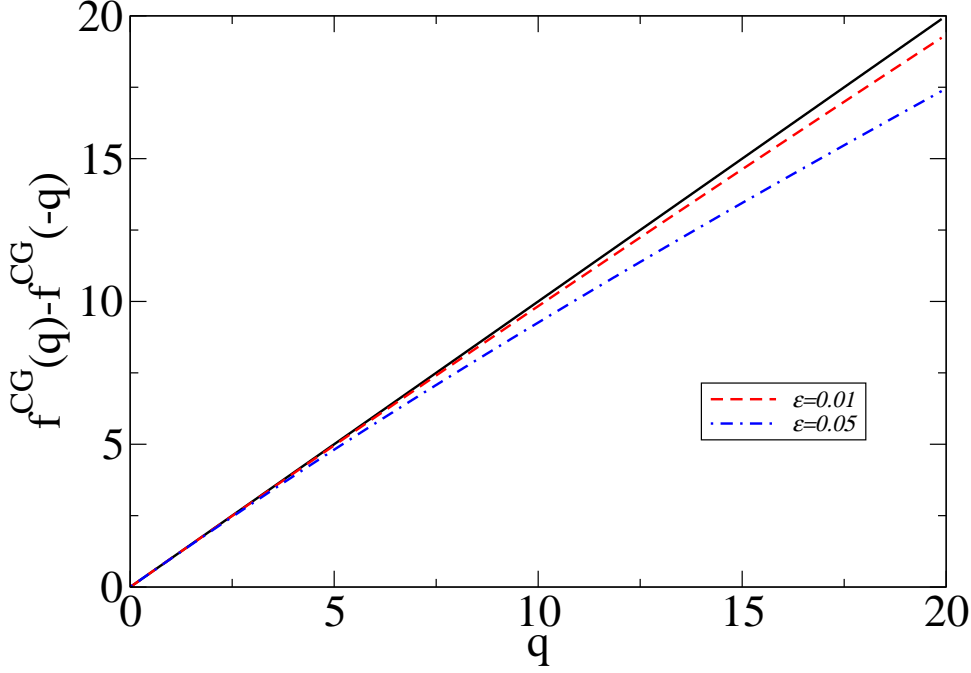


Figure 6. Plot of $f^{\text{CG}}(q) - f^{\text{CG}}(-q)$ as a function of the measured coarse-grained entropy production rate q . The solid line depicts the prediction of the fluctuation relation.

5. Conclusion

In this paper, we have studied stochastic jump processes which can be coarse-grained. The coarse-grained dynamics can be viewed as the dynamics of the system measured using a low resolution apparatus. A perturbation theory, based on a separation of time scales, was used to define a reduced system. The dynamics of this reduced system, which is Markovian (and therefore satisfies an exact fluctuation relation), approximates the coarse-grained system. This reduced system then motivates the definition of a coarse-grained entropy for the coarse-grained dynamics. This entropy was found to approximately satisfy a fluctuation relation. Deviations from this relation are more pronounced for large coarse-grained entropy production rates. This deviation can be interpreted as a result of the existence of internal degrees of freedom, which lead to deviations from Markovian behavior of the coarse-grained entropy.

Let us consider an experiment whose goal is to measure fluctuation relations. The considerations in this paper may be useful in the interpretation of the results of such an experiment. If the results exhibit deviations from the fluctuation relation one may suspect the existence of unobserved degrees of freedom. (However, other mechanisms, leading to deviations from the expected linear behavior, exist [22].) Note that, even if there are internal degrees of freedom, it is entirely possible that the entropy production rates, needed for observing deviations from the fluctuation relation, are so rare that such events will not be measured during the experiment.

The results found in this paper lead to several new questions. For instance, a better understanding of the deviation of the coarse-grained entropy production from

the fluctuation relation is needed. The definition of the coarse-grained entropy used in this paper was motivated by the definition applied to microscopic systems. One may wonder whether it is possible to define another coarse-grained entropy, using only quantities measured by the low resolution apparatus, which will *exactly* satisfy the fluctuation relation. It is also of interest to extend the considerations, obtained here for stochastic systems, to deterministic systems. Finally, we have seen that the existence of a coarse-grained counterpart leads to a physically motivated coarse-grained entropy exhibiting a fluctuation relation. One may wonder whether the reverse is also true. Does the existence of a quantity, which has some physical interpretation, and satisfies a non-trivial fluctuation relation, suggest that the system can be simplified in some way? Such questions are left for future work.

Acknowledgments

We are grateful for start-up funds provided by the University of Maryland.

References

- [1] Chaikin P M and Lubensky T C 1995 *Principles of condensed matter physics* (Cambridge University Press, Cambridge)
- [2] Amit D J 1978 *Field theory, the renormalization group and critical phenomena* (McGraw-Hill Inc., London)
- [3] Fisher M E 1998 *Rev. Mod. Phys.* **70** 653
- [4] Derrida B 2007 *J. Stat. Mech.* **2007** P07023
- [5] Evans D J, Cohen E G D, and Morriss G P 1993 *Phys. Rev. Lett.* **71** 2401
- [6] Evans D J and Searles D J 1994 *Phys. Rev. E* **50** 1645
- [7] Gallavotti G and Cohen E G D 1995 *Phys. Rev. Lett.* **74** 2694
- [8] Bustamante C, Liphardt J, and Ritort F, *Physics Today* **58**, 43 (2005)
- [9] Lebowitz J L and Spohn H 1999 *J. Stat. Phys.* **95** 333
- [10] Kurchan J 1998 *J. Phys. A: Math. Gen.* **31** 3719
- [11] Harris R J and Schütz G M 2007 *J. Stat. Mech.* **2007** P07020
- [12] Zamponi F, Bonetto F, Cugliandolo L F, and Kurchan J 2005 *J. Stat. Mech.* **2005** P09013
- [13] Ilg P and Barrat J L 2006 *J. Phys. Conf. Ser.* **40** 76
- [14] Van Kampen N G 1992 *Stochastic processes in physics and chemistry* (North-Holland, Amsterdam)
- [15] Zia R K P and Schmittmann B, 2006 *J. Phys. A: Math. Gen.* **39** L407
- [16] Snyman I and Geyer H B 2006 *Phys. Rev. B* **73** 144516
- [17] Seifert U 2005 *Phys. Rev. Lett.* **95** 040602
- [18] Imparato A and Peliti L 2007 *J. Stat. Mech.* **2007** L02001
- [19] Gaspard P 2004 *J. Chem. Phys.* **120** 8898
- [20] Schmiedl T and Seifert U 2007 *J. Chem. Phys.* **126** 044101
- [21] Schnakenberg J, 1976 *Rev. Mod. Phys.* **48** 571
- [22] Cohen E G D and van Zon R 2007 *Compt. Ren. Phys.* **8** 507